Computer Science Technical Reports

Computer Science

11-2018

On Efficiency of AS Paths from Users to Content Servers: A Case Study of Netflix

Jinu Susan Kabala
Iowa State University, jsusan@iastate.edu

Lu Ruan
Iowa State University, ruanlu@iastate.edu

On Efficiency of AS Paths from Users to Content Servers: A Case Study of Netflix

Abstract
The majority of the Internet traffic today is content delivery traffic. The performance of content delivery depends on the efficiency of the routing paths from users to content servers and from content servers back to the users. While content providers can control the paths from their servers to the users, they have no control over the paths from users to their content servers and the efficiency of such paths is generally unknown. In this work, we conduct a case study of Netflix to understand the efficiency of the AS paths from various access ISPs to Netflix servers deployed at IXPs in different regions of the world. We discover inefficient AS paths in Europe, North America, and South America. Paths in South America are especially inefficient as many of them leave the continent. We also analyze long paths in each region, explore their causes, and propose ways to avoid long paths.

Keywords
Content delivery networks, Netflix, Internet routing, AS path efficiency

Disciplines
OS and Networks

On Efficiency of AS Paths from Users to Content Servers: A Case Study of Netflix

Jinu Susan Kabala and Lu Ruan
Department of Computer Science
Iowa State University

Abstract

The majority of the Internet traffic today is content delivery traffic. The performance of content delivery depends on the efficiency of the routing paths from users to content servers and from content servers back to the users. While content providers can control the paths from their servers to the users, they have no control over the paths from users to their content servers and the efficiency of such paths is generally unknown. In this work, we conduct a case study of Netflix to understand the efficiency of the AS paths from various access ISPs to Netflix servers deployed at IXPs in different regions of the world. We discover inefficient AS paths in Europe, North America, and South America. Paths in South America are especially inefficient as many of them leave the continent. We also analyze long paths in each region, explore their causes, and propose ways to avoid long paths.

1 Introduction

The majority of the Internet traffic today is content delivery traffic sourced by content delivery networks (CDNs) of large content providers (e.g., Google, Netflix, Facebook) and commercial CDNs (e.g., Akamai, Limelight) [1]. The Internet traffic due to CDNs was 52% of the overall global traffic in 2016, and will cross 71% of overall global traffic by 2021 [2]. The growth of CDNs has resulted in the flattening of the Internet topology [3, 1] as large content providers are relying less on Tier-1 ISPs and peering with larger numbers of lower tier ISPs.

To efficiently deliver content, large content providers deploy servers at IXPs around the world and peer with ISPs at IXPs to bring content close to the users and localize content delivery traffic. However, serving content from a nearby server does not guarantee good performance since latency depends on the *network distance* between a user and the content server instead of the geographical distance. That is, the forward and/or reverse AS path connecting a user and a content server may be particularly long, causing high latency even though the user and the server are close to each other. Krishnan et al. [4] analyze latencies from Google's CDN servers to clients all over the world and observe that many clients experience poor latencies due to long AS paths from Google servers to clients. Such inefficient paths are due to BGP, which is designed to choose policy compliant AS paths and lacks the ability to select AS paths based on performance. To overcome the limitation of BGP, Facebook has developed Edge Fabric [5] and Google has developed Espresso [6]. Both are SDN-based systems providing centralized control of egress routing.

While content providers have control over the AS paths from their servers to users, they do not have control of the AS paths from users to their servers. It is important to understand the efficiency of such paths as it affects the latency of content retrieval. In this paper, we answer the question "*How efficient are AS paths from users to content servers?*" through a detailed case study of Netflix. Netflix is one of the major content delivery players with over 130 million subscribers globally, and its streaming video accounts for 15% of the global Internet traffic. Netflix previously used three CDNs (i.e., Akamai, Limelight, and Level 3) to deliver video content to its subscribers [7]. In 2011, it started to build its own CDN named Open Connect. Open Connet consists of thousands of server appliances deployed within IXPs as well as inside ISP networks to efficiently deliver video content to users [8]. A recent work [9] studies the infrastructure deployment of Netflix and discovers over 8,000 Netflix servers, of which 51% are deployed within IXPs and 49% are deployed inside ISPs. When a Netflix user is served by an IXP server, the quality of experience depends on the efficiency of the AS path connecting the user's access ISP and the Netflix server. In this work, we seek to answer the following questions: Are there inefficient AS paths connecting access ISPs and Netflix IXP servers? How frequent do we observe inefficient AS paths? How does the efficiency of AS paths in different parts of the world differ? To answer these questions, we identify the locations of Netflix IXP servers deployed all over the world, collect AS paths connecting access ISPs to Netflix IXP servers in different regions of the world, and identify inefficient AS paths in each region. We find that Netflix has IXP servers deployed in all five regions of the world: North America (NA), South America, Europe (EU), Asia Pacific (AP), and Africa (AF). We observe inefficient AS paths in Europe, North America, and South America, with Europe having the smallest fraction of inefficient paths and South America has the largest fraction of inefficient paths. We also identify the causes of long AS paths and propose ways to avoid them.

The rest of the paper is organized as follows. Section 2 describes the data used in this study. In section 3, we identify the locations of the Netflix IXP servers. In Section 4, we study the efficiency of AS paths from access ISPs to Netflix IXP servers and investigate the causes of long AS paths. We propose methods to avoid long AS paths in section 5 and conclude the paper in Section 6.

2 Data

Route Views [10] and RIPE NCC [11] have route collectors deployed at IXPs around the world to collect BGP routing tables from various IXP participants, which we refer to as vantage points (VPs). We download routing table snapshots taken by all IXP route collectors on April 11, 2018 at 8am. From the snapshots we extract all routes to Netflix prefixes that are sourced by VPs of type access ISP [1]. We are interested in such routes as they connect end users to Netflix IXP servers. We create a record with three pieces of information from each route: prefix, AS path, and VP location. For example, from a route in RIPE collector rrc10's routing table snapshot we create a record (37.77.188.0/24, 15605 3356 2906, Milan, Italy), which means that VP 15605 located at Milan, Italy uses AS path 15605 3356 2906 to reach Netflix prefix 37.77.188.0/24. The VP location is Milan, Italy since rrc10 is located at MIX in Milan, Italy. Our dataset contains VPs located in all five regions of the world. The first row in Table 1 shows the number of VPs in each region.

Table 1: Number of VPs, prefixes, and paths in each region

Region	**EU**	**NA**	**SA**	**AP**	**AF**
VPs	27	5	12	4	5
Prefixes	32	42	9	12	2
Paths	472	124	70	20	10

3 Identifying Locations of Netflix IXP Servers

The prefixes announced by Netflix (AS2906) correspond to Netflix server clusters deployed at IXPs. In this section we describe how we identify the locations of Netflix prefixes based on the naming convention of Netflix servers.

Netflix includes IATA 3-letter airport codes in the names of its IXP servers to encode the server location [9]. For example, server name *ipv4_1.lagg0.c001.jnb001.ix.nflxvideo.net* contains airport code *jnb*, indicating the server is located at an IXP in Johannesburg, South Africa. To identify the locations of Netflix IXP servers, we first obtain all /24 prefixes announced by Netflix in our dataset. We then use *nslookup* to find the DNS names of the IP addresses in each prefix and use the airport code in the DNS names to obtain the location of each prefix. We have verified that all IP addresses in a /24 prefix contain the same airport code. This means that all servers in a /24 prefix are in the same location.

Our dataset contains a total of 103 /24 Netflix prefixes, of which we are able to obtain the locations for 97 prefixes. The second row in Table 1 shows the number of prefixes located in each region. We see that Netflix has prefixes in all 5 regions of the world. It has the largest number of prefixes in North America, followed by Europe. The prefixes in North America are distributed in 15 cities

[1] We obtain AS types using PeeringDB [12].

in US. The prefixes in Europe are distributed in 14 different countries. In Asia Pacific, Netflix has prefixes in Japan, Hong Kong, Singapore, the United Arab Emirates, Australia, and New Zealand. In South America, Netflix prefixes are found in Brazil and Argentina. In Africa, Netflix has two prefixes, both are located in South Africa.

4 Analysis of AS Paths from Access ISPs to Netflix Prefixes

Netflix deploy IXP servers in a region to deliver video to users in that region. Thus, it is important to understand the efficiency of AS paths connecting users to Netflix IXP servers in the same region. In this section, we analyze the AS paths from access ISPs to Netflix prefixes in each region of the world to understand the efficiency of the AS paths connecting users to Netflix servers in that region. We say an AS path is in region X if the VP (i.e., source access ISP) and the destination Netflix prefix are both located in region X. The third row in Table 1 shows the number of AS paths in each region.

We distinguish two types of paths in each region: *intra-city paths* and *inter-city paths*. An *intra-city path* is an AS path connecting a VP and a Netflix prefix located in the same city. We consider an intra-city path to be efficient if it is 1-hop long (i.e., direct path) and inefficient otherwise. We consider an intra-city path to be long and highly inefficient if it is 3-hop or longer (i.e., the path contains 2 or more intermediate ASes). An *inter-city path* is an AS path connecting a VP and a Netflix prefix located in different cities. We consider an inter-city path to be efficient if it is 1- or 2-hop long and inefficient otherwise. We consider an inter-city path to be long and highly inefficient if it is 4-hop or longer (i.e., the path contains 3 or more intermediate ASes).

In the following, we present our analysis on intra-city paths, inter-city paths, and long AS paths in each region. Our goal is to understand the efficiency of AS paths in each region and identify the causes of long AS paths.

4.1 Intra-City Paths

Table 2 shows the results of intra-city paths in each region, including the total number of intra-city paths, the percentage of direct paths, the percentage of inefficient paths, and the percentage of long paths.

Table 2: Intra-city paths in each region

Region	**EU**	**NA**	**SA**	**AP**	**AF**
Total paths	51	14	38	6	10
Direct paths	98.0%	64.3%	78.9%	100%	100%
Inefficient paths	2.0%	35.7%	21.1%	0%	0%
Long paths	0%	14.3%	10.5%	0%	0%

We see that 98% of the intra-city paths in EU are efficient (i.e., direct) and there are no long paths in EU. On the other hand, 35.7%/21.1% of the intra-city paths in NA/SA are inefficient (i.e., indirect). We have verified that every indirect intra-city path in NA and SA actually connects a VP and a Netflix prefix co-located at the same IXP because Netflix is present at the IXP where the VP is located. These paths are obviously inefficient since the VP reaches a co-located Netflix prefix via one or more intermediate ASes instead of directly. Furthermore, 14.3% of the intra-city paths in NA and 10.5% of the intra-city paths in SA are long (i.e, going though 2 or more intermediate ASes), which are highly inefficient.

We only have 6 intra-city paths in AP, 4 of them are direct and 2 of them are 2-hop long. However, we consider the 2-hop paths to be direct because one path uses AS7606 (IX Australia) and the other path uses AS24115 (Equinix Internet Exchange) as the intermediate AS. This means that the VPs use IXP route servers to peer with Netflix. In fact, IXPs often provide route servers for participants to exchange routing information with multiple peers so that two networks with open peering policy can establish BGP sessions without a bilateral peering agreement in place.

There are 10 intra-city paths in AF. All of these paths have the VP and the Netflix prefix co-located at NAPAfrica Johannesburg, South Africa, and all the paths are direct. This indicates that the VPs peer with Netflix at NAPAfrica Johannesburg to directly reach Netflix prefixes. Gupta et al. [13] observe in 2014 that 66.8% of paths from routers in home networks to Google cache servers in Africa left the continent, and ISPs in Africa were peering at large IXPs in Europe instead of peering at local IXPs. The authors note that as more cache nodes are deployed in the region and more traffic could remain local, the peering ecosystem may rapidly evolve to include more local peering links. Our results confirm that this is indeed happening: Netflix is deploying servers at IXPs in Africa, and African ISPs are peering with Netflix at local IXPs to reach Netflix prefixes directly.

4.2 Inter-City Paths

Table 3 shows the results of inter-city paths in each region, including the total number of inter-city paths, the percentage of direct paths, the percentage of inefficient paths, and the percentage of long paths. No results are shown for Africa because our dataset does not contain any inter-city paths in AF.

Table 3: Inter-city paths in each region

Region	**EU**	**NA**	**SA**	**AP**
Total paths	421	110	32	14
Direct paths	24.9%	12.7%	12.5%	100%
Inefficient paths	15.7%	27.3%	43.8%	0%
Long paths	3.3%	9.1%	43.8%	0%

There are 14 inter-city paths in AP, all of which are direct. In fact, some of the inter-city paths in AP are 2-hop long, but are considered to be direct because they contain an IXP as an intermediate AS. Thus, all Asia Pacific VPs in our dataset peer with Netflix either directly or through an IXP route server to reach Netflix prefixes located in other cities in Asia Pacific. We note that an inter-city path can be direct because an VP at IXP X in one city can reach a Netflix prefix at IXP Y in another city in one hop if the VP is also present at Y.

We now look at EU, NA, and SA, all of which have inefficient as well as long inter-city paths. Among the three regions, EU has the lowest percentage of inefficient paths (15.7%) and the lowest percentage of long paths (3.3%). SA has the highest percentage (43.8%) of inefficient paths, and all of them are long. In NA, 27.3% of the paths are inefficient, and one third of them are long.

In summary, we observe inefficient intra-city and inter-city paths in EU, NA, and SA, with EU having the best path efficiency and SA having the worst path efficiency. We do not observe any inefficient paths in AP and AF. This does not imply that all paths in AP and AF are efficient because our dataset contains a limited number of VPs which give us limited visibility of AS paths in these two regions.

4.3 Long Paths

We expect the intermediate ASes in an AS path to be transit ISPs. Thus, we consider an AS path to be abnormal if it contains an intermediate AS that is not a transit ISP. In this section, we identify abnormal long paths and explore their causes.

4.3.1 Long Intra-City Paths

An intra-city path is considered long if it is 3-hop or longer. Our dataset contains a total of 6 long intra-city paths, 2 of which are abnormal. Both of the abnormal paths are in North America and go through an access ISP immediately before reaching AS2906: one path goes through AS7922 (Comcast), while the other goes through AS11492 (Cable One). It is surprising to see an access ISP acting as an upstream provider for Netflix. A possible explanation is that AS2906 announces certain prefixes to AS7922/AS11492 so that customers of AS7922/AS11492 can directly reach those prefixes. However, AS7922/AS11492 incorrectly announced the prefixes to their neighbors so that other ASes may reach the Netflix prefixes via AS7922/AS11492. Such incorrect announcement results in long AS paths that reach Netflix prefixes through AS7922 and AS11492, causing them to act as transit providers for Netflix.

4.3.2 Long Inter-City Paths

An inter-city path is considered long if it is 4-hop or longer. Our dataset contains a total of 38 long inter-city paths, 21 of which are abnormal. Some of the abnormal paths contain an access ISP, as in the case of intra-city paths. Specifically, some paths in Europe go though AS3303 (Swisscom) and some paths in North America go through AS11492 (Cable One). It is likely that AS3303 and AS11492 incorrectly announced Netflix prefixes to their neighbors, causing other ASes to reach Netflix prefixes via them. There are 14 abnormal long paths in South America, all of which go through a cloud provider. Specifically, two paths go through AS16685 (TIVIT) and the other 12 paths go through AS27988 (SYTSA). Again, it is likely that AS16685 and AS27988 mistakenly announced Netflix prefixes to their neighbors even though they are not supposed to do so.

We now take a closer look at the 12 paths that go through AS27988. One of them is sourced by AS263047, which reaches prefix 23.246.52.0/24 (located in Buenos Aires, Argentina) using AS path 263047 6939 27988 52376 2906. AS263047 has a public looking glass server located in Rio de Janeiro, Brazil. We run MTR on the looking glass server to trace a route to an IP address in prefix 23.246.52.0/24. The trace verifies that the AS path traversed is 263047 6939 27988 52376 2906. The average RTT is 242.6ms, which is very high considering the distance between the source and the destination is around 1200 mile. One of the routers on the traceroute path is *10ge9-16.core1.mia1.he.net*, which belongs to AS6939 (Hurricane Electric) and is located in Miami, FL as the router name contains airport code '*mia*'. This means that packets sent from a server located in Brazil and destined to a Netflix prefix located in Argentina first leaves South America to go to Miami in US and then comes back to South America. The reason that AS6939 hands off traffic to AS27988 at Miami is that FL-IX at Miami is the only common IXP where both 6939 and 27988 are present. The other 11 paths contain either AS6939 or AS22356 as the hop immediately preceding AS27988. From PeeringDB we see that AS6939 and AS22356 both share a single common IXP (i.e., FL-IX at Miami) with AS27988. Thus, those 11 paths also leave South America by going through FL-IX at Miami. Such circuitous paths are highly inefficient and result in high RTT.

In summary, we observe abnormal long paths in EU, NA, and SA. These paths contains an access ISP or a cloud provider as an intermediate AS, which is likely caused by incorrect announcements of Netflix prefixes by the access ISP or cloud provider. In SA, we observe many highly inefficient paths that leave the continent to detour through US because two intermediate ASes on the path only peer in US.

5 Avoiding Long Paths

As discussed in the previous section, some paths are long because they contain a non-transit ISP (e.g., an access ISP, a cloud provider), which announces a prefix to a neighbor even though it is not supposed to do so. These abnormal long paths can be avoided if the operators of the non-transit ISPs carefully review their route announcement policies to eliminate incorrect announcements.

We now consider long AS paths that contain only transit ISPs. Such long paths arise because each individual AS along the path imports/exports routes based on its own routing policy without any cooperation with other ASes. We propose that an access ISP and a content provider could cooperate to avoid long paths by sharing neighbor information. Specifically, an access ISP and a content provider can exchange their lists of neighbors. If the two ASes have a common neighbor, then that neighbor can be used to create a 2-hop path between them. If there is no common neighbor, the access ISP and the content provider can each choose a neighbor so that the two neighbors are directly connected [2]. This will result in a 3-hop path between the access ISP and the content provider. Since transit ISPs in today's Internet are densely interconnected, it is highly likely that information sharing between access ISPs and content providers could avoid paths that are longer than 3 hops. Furthermore, when a 3-hop path is chosen, access ISPs and content providers should choose neighbors that interconnect within the region to avoid circuitous paths that leave the region.

It has been shown in [14] that there is significant cooperation between adjacent ISPs to avoid particularly poor routes or load-balance traffic across multiple peering links. We argue that such cooperation is needed between access ISPs and content providers in order for them to select efficient AS paths to reach each other's prefixes.

6 Conclusion

We have conducted a detailed case study of Netflix to understand the efficiency of AS paths from users to Netflix IXP servers. We identify the locations of Netflix prefixes, and analyze the AS paths from access ISPs to Netflix prefixes in each region of the world. We observe inefficient AS paths in Europe, North America, and South America. While most of the paths in Europe are efficient, both North America and South America have a notable percentage of inefficient intra-city and inter-city paths. South America has the worst path efficiency: 43.8% of the inter-city paths are as long as 4 hops and most of them leave the continent to detour through US. We find that some of the long paths are abnormal in that they go though a non-transit ISP such as an access ISP or a cloud provider. These abnormal paths are likely caused by the non-transit ISP incorrectly announcing a Netflix prefix to its neighbors, causing other ASes to reach Netflix prefixes through it.

We propose that access ISPs and content providers can avoid using long AS paths to reach each other's prefixes by sharing neighbor information. We note that this is a short term solution that requires direct communication between access ISPs and content networks. A long term solution would be to enhance BGP to facilitate information sharing between ASes so that any AS can choose an efficient path to reach any prefix. We leave this as a future work.

[2] One can determine if two ASes are directly connected using BGP data provided by Route Views and RIPE NCC.

References

[1] C. Labovitz, S. Iekel-Johnson, D. McPherson, J. Oberheide, and F. Jahanian, "Internet Inter-domain Traffic," in *Proceedings of SIGCOMM*, 2010, pp. 75–86.

[2] The Zettabyte Era: Trends and Analysis. [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/vni-hyperconnectivity-wp.html

[3] P. Gill, M. Arlitt, Z. Li, and A. Mahanti, "The Flattening Internet Topology: Natural Evolution, Unsightly Barnacles or Contrived Collapse?" in *Proceedings of PAM*, 2008, pp. 1–10.

[4] R. Krishnan, H. V. Madhyastha, S. Srinivasan, S. Jain, A. Krishnamurthy, T. Anderson, and J. Gao, "Moving Beyond End-to-end Path Information to Optimize CDN Performance," in *Proceedings of IMC*, 2009, pp. 190–201.

[5] B. Schlinker, H. Kim, T. Cui, E. Katz-Bassett, H. V. Madhyastha, I. Cunha, J. Quinn, S. Hasan, P. Lapukhov, and H. Zeng, "Engineering Egress with Edge Fabric: Steering Oceans of Content to the World," in *Proceedings of SIGCOMM*, 2017, pp. 418–431.

[6] K.-K. Yap, M. Motiwala, J. Rahe, S. Padgett, M. Holliman, G. Baldus, M. Hines, T. Kim, A. Narayanan, A. Jain, V. Lin, C. Rice, B. Rogan, A. Singh, B. Tanaka, M. Verma, P. Sood, M. Tariq, M. Tierney, D. Trumic, V. Valancius, C. Ying, M. Kallahalla, B. Koley, and A. Vahdat, "Taking the Edge off with Espresso: Scale, Reliability and Programmability for Global Internet Peering," in *Proceedings of SIGCOMM*, 2017, pp. 432–445.

[7] V. K. Adhikari, Y. Guo, F. Hao, M. Varvello, V. Hilt, M. Steiner, and Z. L. Zhang, "Unreeling Netflix: Understanding and Improving Multi-CDN Movie Delivery," in *Proceedings of IEEE INFOCOM*, 2012, pp. 1620–1628.

[8] Netflix Open Connect. [Online]. Available: https://openconnect.netflix.com/en/

[9] T. Böttger, F. Cuadrado, G. Tyson, I. Castro, and S. Uhlig, "Open Connect Everywhere: A Glimpse at the Internet Ecosystem Through the Lens of the Netflix CDN," *SIGCOMM Comput. Commun. Rev.*, vol. 48, no. 1, pp. 28–34, Apr. 2018.

[10] Route Views. [Online]. Available: http://www.routeviews.org/

[11] RIPE NCC Routing Information Service. [Online]. Available: https://www.ripe.net/analyse/internet-measurements/routing-information-service-ris

[12] PeeringDB. [Online]. Available: https://www.peeringdb.com/

[13] A. Gupta, M. Calder, N. Feamster, M. Chetty, E. Calandro, and E. Katz-Bassett, "Peering at the Internet's Frontier: A First Look at ISP Interconnectivity in Africa," in *Proceedings of PAM*, 2014, pp. 204–213.

[14] N. Spring, R. Mahajan, and T. Anderson, "Quantifying the Causes of Path Inflation," in *Proceedings of SIGCOMM*, 2003, pp. 113–124.